

# Improved Achievable Information Rates by Optimized Four-Dimensional Demappers in Optical Transmission Experiments

Tobias Fehenberger<sup>(1)</sup>, Tobias A. Eriksson<sup>(2)</sup>, Alex Alvarado<sup>(3)</sup>, Magnus Karlsson<sup>(2)</sup>, Erik Agrell<sup>(4)</sup>, and Norbert Hanik<sup>(1)</sup>

<sup>(1)</sup>*Institute for Communications Engineering, Technische Universität München (TUM), 80333 Munich, Germany*

<sup>(2)</sup>*Dept. of Microtechnology and Nanoscience, Chalmers University of Technology, 41296 Gothenburg, Sweden*

<sup>(3)</sup>*Optical Networks Group, Dept. of Electronic & Electrical Engineering, UCL, London, WC1E 7JE, UK*

<sup>(4)</sup>*Dept. of Signals and Systems, Chalmers University of Technology, 41296 Gothenburg, Sweden*

*tobias.fehenberger@tum.de*

**Abstract:** We experimentally study different four-dimensional demappers in a dispersion-managed fiber system. The proposed blind algorithm is shown to offer gains of 0.2 bits per 4D symbol for DP-16QAM.

**OCIS codes:** (060.2330) Fiber optics communications, (060.4080) Modulation.

## 1. Introduction

Since the advent of coherent receivers with digital signal processing (DSP), the combination of high-order modulation formats, in particular quadrature amplitude modulation (QAM), with advanced forward error correction (FEC) has enabled increased data rates [1]. Bit-wise decoders for these FEC schemes have become the de-facto standard, mainly due to their excellent trade-off between performance and complexity [2]. For these decoders, the relevant channel output are the log-likelihood ratios (LLRs), which are calculated from the noisy symbols after DSP.

The calculation of LLRs and the estimation of achievable information rates (AIRs) require an assumption on the channel model. Typically, an additive white Gaussian noise (AWGN) is used for this purpose at the demapper. The statistics of the AWGN channel is characterized by its channel covariance matrix and mean values. When residual memory after DSP is neglected, at most four dimensions (4D) must be considered in the case of dual-polarization (DP) modulation. Restricting the noise model to circularly symmetric (CS) two-dimensional (2D) statistics is a common choice in optics [3], and has been shown to give identical AIRs, in comparison to more complex AWGN statistics, for long-haul fiber systems without inline dispersion-management (DM) [4]. However, in the presence of uncompensated phase noise, for example due to insufficient carrier phase estimation (CPE) or due to large nonlinear fiber effects, using non-CS statistics instead of a CS Gaussian model improves performance, as shown numerically in [5]. Residual phase noise in the LLR calculation has been studied numerically in [6, 7]. Transmission experiments have shown a decrease in coding gain if the LLR calculation is not properly adapted to the statistics of the DSP output [8]. Additionally, the mean values of the received distributions can be adjusted to correct for transceiver imperfections [9].

In this work, we experimentally study AIRs after transmission of DP-16QAM over a dispersion-managed link. Gains of up to 0.2 bits per 4D symbol are found by using the blind expectation maximization (EM) algorithm to estimate 4D AWGN statistics. To the best of our knowledge, this is the first experimental study of the impact of 4D channel statistics and 4D adaptive mean values on AIRs.

## 2. Achievable Information Rates for Multidimensional Demappers

For a fixed modulation format  $\mathcal{X} = \{\mathbf{s}_1, \dots, \mathbf{s}_M\}$  of order  $M$  and a fixed memoryless channel, mutual information (MI) is the largest AIR. In order to calculate MI, the channel law must be known, which is in general not the case in optical communication systems. Assuming an auxiliary channel  $q_{\mathbf{Y}|\mathbf{X}}$  between the 4D channel input  $\mathbf{X}$  and output  $\mathbf{Y}$  at the demapper gives a lower-bound estimate of the MI, which is an AIR for a receiver that uses  $q_{\mathbf{Y}|\mathbf{X}}$ . For simplicity, we refer to this quantity as MI, emphasizing that it is only a lower-bound estimate. We further note that the better  $q_{\mathbf{Y}|\mathbf{X}}$  resembles the actual channel, the tighter this lower bound becomes and the closer MI is to its true value. MI is computed via Monte-Carlo integration from  $N$  input-output pairs  $(\mathbf{x}_i, \mathbf{y}_i)$  with  $i = 1, \dots, N$  as [10, Sec. IV-B]

$$\text{MI} \approx \frac{1}{N} \sum_{i=1}^N \log_2 \frac{q_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}_i|\mathbf{x}_i)}{q_{\mathbf{Y}}(\mathbf{y}_i)}. \quad (1)$$

If we restrict the decoder to be bit-wise, i.e., when LLRs are calculated prior to binary FEC decoding, the MI in (1) is not achievable. A common AIR for this practically relevant case is the generalized mutual information (GMI) [11],

$$\text{GMI} \approx m - \frac{1}{N} \sum_{k=1}^m \sum_{i=1}^N \log_2(1 + e^{(-1)^{b_{k,i}} \cdot \text{LLR}_{k,i}}), \quad (2)$$

where  $m = \log_2 M$  is the number of bits per symbol,  $k$  denotes the bit position per symbol,  $i$  indices the  $i^{\text{th}}$  received sample, and  $b_{k,i}$  are sent bits. For the LLRs in (2), an auxiliary channel must be assumed, in analogy to the MI calculation in (1). Since the auxiliary channel corresponding to the true optical channel is not known, the LLRs are not matched and, in principle, an optimization over a non-negative parameter is required in (2) [11, Sec. III-C]. This optimization is omitted in this work, possibly resulting in a lower GMI estimate.

For  $q_{\mathbf{Y}|\mathbf{X}}$ ,  $d$ -dimensional multivariate Gaussian distributions are used. The LLR at bit position  $k$  of the  $i^{\text{th}}$  symbol is

$$\text{LLR}_{k,i} = \log \frac{q_{\mathbf{Y}|B_k}(\mathbf{y}_i | b_{k,i}=0)}{q_{\mathbf{Y}|B_k}(\mathbf{y}_i | b_{k,i}=1)} = \log \frac{\sum_{j:\mathbf{s}_j \in \mathcal{X}_0^k} q_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}_i | \mathbf{s}_j)}{\sum_{j:\mathbf{s}_j \in \mathcal{X}_1^k} q_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}_i | \mathbf{s}_j)} = \log \frac{\sum_{j:\mathbf{s}_j \in \mathcal{X}_0^k} \frac{1}{\sqrt{(2\pi)^d |\Sigma_j|}} \exp\left(-\frac{1}{2}(\mathbf{y}_i - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_j)\right)}{\sum_{j:\mathbf{s}_j \in \mathcal{X}_1^k} \frac{1}{\sqrt{(2\pi)^d |\Sigma_j|}} \exp\left(-\frac{1}{2}(\mathbf{y}_i - \boldsymbol{\mu}_j)^T \Sigma_j^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_j)\right)}, \quad (3)$$

where  $B_k$  denotes bit  $k$  of  $\mathbf{X}$ , and  $\mathcal{X}_0^k$  and  $\mathcal{X}_1^k$  are the set of constellations points where the  $k^{\text{th}}$  bit equals to 0 and 1, respectively. Matrix determinant, inverse, and transpose are denoted by  $|\bullet|$ ,  $(\bullet)^{-1}$ , and  $(\bullet)^T$ , respectively. Further,  $\Sigma_j$  is a  $d \times d$  covariance matrix corresponding to the  $j^{\text{th}}$  sent symbol  $\mathbf{s}_j$  for  $j = 1, \dots, M$ , and  $\boldsymbol{\mu}_j$  denotes the mean value that corresponds to  $\mathbf{s}_j$  being sent. We note that both MI in (1) and GMI in (2) depend on the auxiliary channel  $q_{\mathbf{Y}|\mathbf{X}}$ . Thus, the choice of  $\Sigma_j$  and  $\boldsymbol{\mu}_j$  has a direct impact on the AIRs, which is discussed in the following.

We distinguish two cases for obtaining the mean values  $\boldsymbol{\mu}_j$ . Either they are chosen to be identical to the sent constellation, i.e.,  $\boldsymbol{\mu}_j = \mathbf{s}_j$  for all  $j$ , which we call *static* mean values. Alternatively,  $\boldsymbol{\mu}_j$  can be obtained from the received samples  $\mathbf{y}$  conditioned on  $\mathbf{s}_j$  being sent. We refer to this conditional sample means as *adaptive* mean values. In the special case of 2D CS Gaussian noise,  $\Sigma_j$  is the  $2 \times 2$  identity matrix times half the noise variance for all samples. In all other cases, no restrictions are made on  $\Sigma_j$ , i.e., we allow different variances on all dimensions and correlations between them. Estimating the parameters  $\Sigma_j$  and  $\boldsymbol{\mu}_j$  can be done with either a data-aided (DA) method or a blind algorithm. For the DA scheme, a block of symbols at the beginning of the DSP output is used. These samples are disregarded afterwards in order not to overestimate AIRs. The expectation maximization (EM) algorithm [12, Sec. II] is employed as a blind algorithm to find the mean values and covariance matrices of the received samples.

### 3. Experimental Setup

An arbitrary waveform generator running at 60 GSamples/s was used to generate 10 GBaud DP-16QAM signals with a root-raised cosine pulse shape (50% roll-off). The electrical signals and a laser with  $\sim 150$  kHz linewidth were input into two IQ-modulators, generating 7 wavelength division multiplexing (WDM) channels spaced at 15 GHz. Polarization multiplexing is emulated using a split and decorrelation stage. The launch power was set to the optimum value ( $-9$  dBm per WDM channel for all distances).

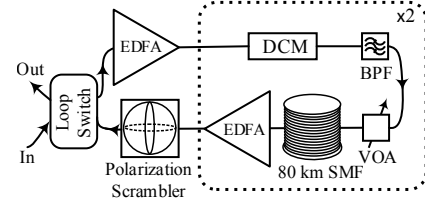


Fig. 1. Setup of the recirculating loop.

The signals were propagated over a recirculating loop, shown in Fig. 1, based on two spans of 80 km standard single-mode fiber (SMF). Before each span, inline dispersion compensation is done with a dispersion-compensating module (DCM) based on fiber-Bragg gratings that inflict no nonlinear distortions. This enables the use of a single Erbium-doped fiber amplifier (EDFA) per span. An optical band-pass filter (BPF) suppresses out-of-band noise and the power to each span is set by a variable optical attenuator (VOA).

After propagation, the center channel was detected using a conventional optical coherent receiver with a free-running  $\sim 150$  kHz laser as local oscillator. The electrical signals were sampled at 50 GSamples/s using a real-time oscilloscope. For polarization-demultiplexing and equalization, the constant-modulus algorithm was used for pre-convergence of the taps, which are further adapted by a decision-directed least-mean-square equalizer. Within this equalizer, CPE was carried out using a blind phase search. With more than  $10^6$  symbols in total, MI and GMI (with Gray labeled symbols) were estimated using parameters obtained by the DA and EM methods outlined in Sec. 2. For all 2D estimates, MI and GMI were calculated per polarization and added up.

### 4. Results

Figures 2 (a) and (b) show MI and GMI, respectively, in bits per 4D symbol (bits/4D-sym) vs. transmission distance. We compare four different scenarios for obtaining the parameters  $\Sigma_j$  and  $\boldsymbol{\mu}_j$ . The conventional method is to use a

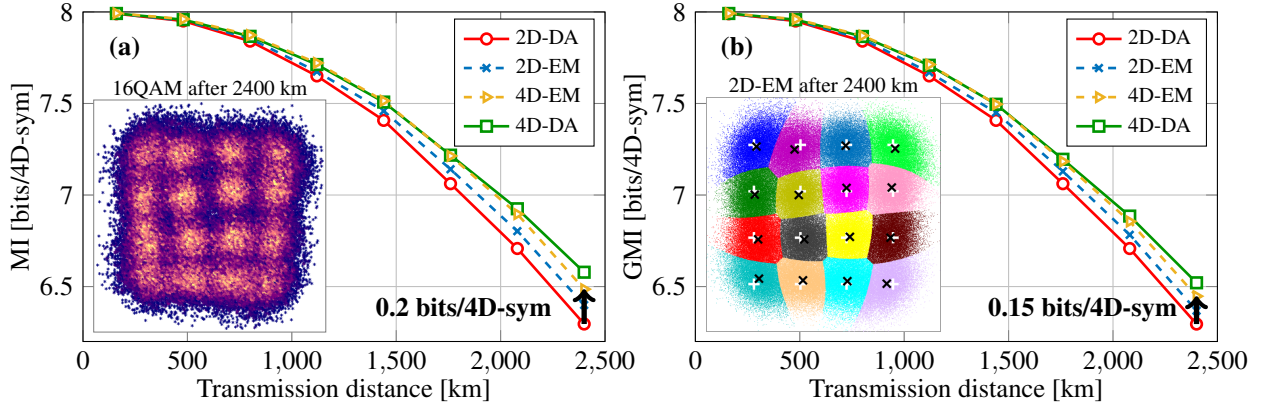


Fig. 2. MI (a) and GMI (b) in bits per 4D symbol vs. transmission distance. Insets: 16QAM received distribution (left) and corresponding 2D-EM output (right) with adaptive mean values (black crosses). Static mean values (white plus signs) are shown as reference.

2D DA estimate with a CS covariance matrix and static mean values (2D-DA). An upper bound is shown by using a 4D Gaussian distribution with adaptive mean values, obtained with the DA approach (4D-DA). The EM algorithm always gives adaptive mean values and covariance matrices, and we distinguish only between the 2D (2D-EM) and 4D (4D-EM) case, which we show in Fig. 2. For DP-16QAM, the 4D-EM output consists of 256 covariance matrices of dimensionality  $4 \times 4$  and 256 mean values in 4D.

We see in Fig. 2 that MI and GMI qualitatively follow the same trend for all four cases (with slightly larger rates in the case of MI), and focus on GMI in the following. No significant differences between the four GMIs curves are observed for short distances because little nonlinear distortion has accumulated yet and the channel statistics are close to CS. For larger distances, gains over 2D-DA are achieved by the EM algorithm. After 2400 km, the GMI gain is 0.05 bits/4D-sym for 2D-EM but increases to 0.15 bits/4D-sym by using 4D-EM. For MI, the gain is 0.2 bits/4D-sym. Both 4D-EM MI and GMI are close to their respective 4D-DA upper bound for all distances, showing the excellent estimation performance of the blind EM algorithm.

In the insets of Fig. 2, the received distribution after 2400 km (left) and the corresponding output of the 2D-EM algorithm (right) are shown. The nonlinear decision regions (colored areas) and the deviation of the adaptive mean values (black crosses) from the static mean values (white plus signs) illustrate how the EM attains AIR improvements over the standard 2D-DA case.

## 5. Conclusions

We have experimentally investigated data-aided and blind estimates of multidimensional Gaussian channel statistics. By employing the blind EM algorithm to perform 4D channel estimation with adaptive mean values, AIR improvements over the standard 2D case are obtained. These gains are practically achievable for 4D demappers in optical communication systems.

## Acknowledgments

Research supported by the EPSRC project UNLOC (EP/J017582/1), UK, and the Swedish Research Council.

## References

1. D. J. Richardson, "Filling the light pipe," *Science* **330**(6002), 2010.
2. A. Alvarado and E. Agrell, "Four-dimensional coded modulation with bit-wise decoders for future optical communications," *JLT* **33**(10), 2015.
3. P. Poggiolini *et al.*, "The GN-model of fiber non-linear propagation and its applications," *JLT* **32**(4), 2014.
4. T. A. Eriksson *et al.*, "Four-dimensional estimates of mutual information in coherent optical communication experiments," *Proc. ECOC, We.4.6.5*, 2015.
5. D. Marsella *et al.*, "Maximum likelihood sequence detection for mitigating nonlinear effects," *JLT* **32**(5), 2014.
6. J. Cho *et al.*, "Analysis of soft-decision FEC on non-AWGN channels," *Opt. Express* **20**(7), 2012.
7. I. B. Djordjevic and T. Wang, "On the LDPC-coded modulation for ultra-high-speed optical transport in the presence of phase noise," *Proc. OFC, OM2B.1*, 2013.
8. Y. Zhao *et al.*, "Beyond 100G optical channel noise modeling for optimized soft-decision FEC performance," *Proc. OFC, OW1H.3*, 2012.
9. D. Millar *et al.*, "Detection of a 1Tb/s superchannel with a single coherent receiver," *Proc. ECOC, Mo.3.3.1*, 2015.
10. M. Secondini *et al.*, "Achievable information rate in nonlinear WDM fiber-optic systems with arbitrary modulation formats and dispersion maps," *JLT* **31**(23), 2013.
11. A. Alvarado *et al.*, "Replacing the soft-decision FEC limit paradigm in the design of optical communication systems," *JLT* **33**(20), 2015.
12. Y. Chen *et al.*, "Fiber nonlinearity compensation for dispersion unmanaged PDM 8-QAM CO-OFDM using expectation maximization," *Proc. OECC, WR4-7*, 2013.